International Journal of Advanced Research in Science, Engineering and Technology

Vol. 13, Issue 1, January 2026

ISSN: 2350-0328

IJARSET IMPACT FACTOR 7.150

Early Detection of Autism Spectrum Disorder in Children Using Machine Learning

Dr.S.Tharani, M.Karthiyayini, M.Priyatharshini, A.Yogeswari, V.Vinodkanna

Assistant Professor, Department of Computer Science, United College of Arts and Science, Coimbatore.
Assistant Professor, Department of Computer Science, United College of Arts and Science, Coimbatore.
Assistant Professor, Department of Computer Science, United College of Arts and Science, Coimbatore.
Assistant Professor, Department of Computer Science, United College of Arts and Science, Coimbatore.
Assistant Professor, Department of Computer Science, United College of Arts and Science, Coimbatore.

ABSTRACT: One might simply define Autism Spectrum Disorders (ASD) as a heterogeneous neurodevelopmental condition affecting communication, social interaction, and learning. The diversity in the clinical presentation of ASD, especially in children, necessitates that early detection and intervention be at the core of the treatment for a more gainful outcome. Here, we explore an integrated multimodal dataset consisting of facial image stimuli, eye-tracking scan path data, and metadata such as age, gender, ASD classification, and clinical severity as measured by the Childhood Autism Rating Scale (CARS). The eye-tracking data were collected in two-dimensional sequential coordinate form, which were afterwards mapped onto facial images to visualize fixation and saccade movements. The dataset was divided into ASD and Non-ASD groups for behavioral comparison purposes. The preprocessing included gaze trail rectification, temporal alignment, and feature extraction. ASD severity prediction was made using a classification approach via Random Forest, Logistic Regression, and SVM. Behavioral findings show that children with ASD have fewer fixations on the face and even show irregular scanning paths, thus depicting the importance of combining eye-tracking and interpretable machine learning for early and personalized screening and intervention.

KEYWORDS: Machine Learning,, Logistic Regression, Support Vector Machine (SVM), Autism Spectrum Disorder (ASD),Random Forest, Eye-Tracking ASD.

I. INTRODUCTION

Around the world, Autism Spectrum Disorder (ASD) is growing in prevalence, and it is thereby creating the need for early detection and intervention. Traditional diagnostic methods are assessments dependent on the subjective judgment of the assessment. So it becomes possible that children could experience delays in support. It can be seen that there is now an increasing demand for objective, scalable tools that will actually improve diagnosis. Eye-tracking can non-invasively capture behavioral markers that are linked to the ASDs. Applying machine learning to it opens up transformative possibilities for the screening, interpretation, and management of ASDs.

A. Understanding Autism Spectrum Disorder and Diagnostic Challenges

ASD is a complex neurodevelopmental disorder that manifests itself through enduring differences in social communication, interaction, and behavior. Although symptoms usually start during early childhood, they differ from one individual to another, making ASD highly heterogeneous and difficult to diagnose. Early diagnosis is essential, and it allows for timely interventions that result in major improvements in developmental outcomes. On the contrary, traditional general methods of diagnosis-the clinical observations and behavioral assessments- are usually long, subjective, and unavailable in many places. Future dependency will be compared with scalable, objective technology-assisted screening solutions.

B. Study Objective and Significance

The main goal of this study is the development of a lightweight, interpretable, reproducible classification pipeline for detecting ASD traits and estimating severity from image-derived features. The feasibility of early autism detection using simple scanpath statistics and demographic metadata, without the need for highly specialized hardware or complex imaging, is emphasized. This development could greatly enhance clinical diagnostic support while also fostering the emergence of scalable digital screening tools, particularly in under-resourced settings. Machine learning combined with eye-tracking data thus furthers the efforts toward creating early, accessible, and data-driven autism diagnostics.

II. RELATED WORK

Liu et al. (2016) analyzed the face-scanning patterns and trained SVMs on the corresponding eye movement data. They classified subjects with 88.5% accuracy as either ASD or TD children. Particularly useful was looking at the regions of the eyes and mouth.

Tsuchiya et al., (2021) developed application of Gazefinder for recording visual attention in school-aged children. They applied the best-fit age-based model to yield classification accuracies of approximately 78%.It has shown that age-sensitive models are feasible for real- field screening.

Wan et al. (2019) employed Support Vector Machines (SVM) to categorize ASD and TD groups using fixation durations. More than 80% accuracy was attained. This work stemmed from a motivational argument emphasizing a temporal gaze metric toward ASD diagnosis.

Kanhirakadavath & Chandran (2022) constructed motion based gaze features for deep neural networks. They reported an increase in model accuracy by 23% using velocity and jerk metrics. This study is based on dynamic eye behavior and not static fixations.

Aneva et al.(2018) were eye-tracking study for use during web interactions in adolescents with ASD. More or less accurate up to 75% based upon logistic regression for class separation. This was proved that gaze can reflect cognitive patterns beyond the stimuli face.

III. EXISTING SYSTEM

It has been noted that traditional behavior-based assessments of ASD are largely subjective, so the best practices for early diagnosis are attracting attention in psychology and neuroscience and AI. Eye-tracking technology together with machine learning makes it possible to objectively study gaze behaviors in which children with ASD are less attentive to faces and social cues. Models like Random Forest and SVM have good performance, while the challenges include small, non-diverse datasets, a lack of explainability, and over-reliance on gaze features. Using metadata like age, gender, and CARS scores might help increase diagnostic accuracy and aid in estimating severity. A multimodal explainable framework would not only strengthen the prediction but would also help the clinicians by giving clearer insight. Such advancements could increase the reliability, scalability, and accessibility of screening for ASD in service of early intervention.

IV. METHODOLOGY

This study is intended to establish an efficient machine learning pipeline for classifying children showing signs of Autism Spectrum Disorder as compared to those not showing signs of ASD by means of eye-tracking scanpath images. The approach consists of ways to carry out image preprocessing, statistical feature extraction, supervised learning algorithms, and model evaluation, thus making it very helpful for the early diagnosis of the conditions.

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 13, Issue 1, January 2026

ISSN: 2350-0328

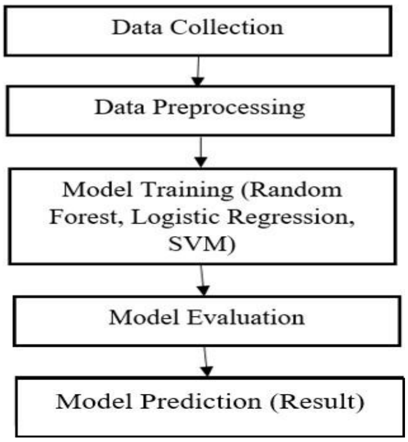

Fig 1. Workflow of the Proposed System

V. PROPOSED ALGORITHM

Three ensemble and supervised classifiers were trained to predict Autism Spectrum Disorder (ASD) in children based on extracted scanpath image features:

- **Random Forest:** To predict whether a child has ASD or not, the random forest, which is a multitude of decision trees, was used to analyze the scanpath features plus some demographic variables such as age and gender. Because it addresses imbalanced data, minimizes overfitting, and achieves high accuracy, it goes well with early, non-invasive screening for autism.
- **Logistic Regression:** A simple, easily interpretable baseline, it maps its features to probability of ASD and non-ASD, which helps understand about feature impact.
- **Support Vector Machine:** It provides optimal separating boundaries, while being very effective on small datasets. Complex non-linear patterns are modeled by it, and that helps capture subtle differences in gaze behavior.

The data is split into training and test data at an 80:20 ratio, and the models were assessed according to four major classification measures, namely: accuracy, precision, recall, and F1- score. Among all the models trained, Random Forest classifier gave the optimized performance and hence will be the one that will be used for final prediction of ASD, keeping balance between accuracy and interpretability.

VI. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed predictive system effectively differentiates ASD from non-ASD children using minimally invasive visual and demographic features. The models demonstrated both strong accuracy and interpretability, supporting early screening and aiding clinical decision-making. These results highlight the system's potential to translate predictive analytics into actionable insights for autism awareness and early intervention.

A. Model Performance Evaluation

Random Forest, Logistic Regression, and SVM are the models used to train on eye-tracking features using an 80:20 split for training and testing, respectively, evaluated using standard metrics. Random Forest was the best-performing algorithm (87.19% training, 78.18% test accuracy with strong precision/recall), followed by Logistic Regression and SVM with 77% test accuracy and a more balanced generalization. Overall, these models were effective in the classification of ASD, with Random Forest being better suited for application in early screening.

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 13, Issue 1, January 2026

ISSN: 2350-0328

IJARSET IMPACT FACTOR 7.150

Models	Class	Precision	Recall	F1-Score	Accuracy
Random Forest	0	0.93	0.89	0.84	78.18 %
	1	0.79	0.89	0.84	
Logistic Regression	0	0.78	0.86	0.82	77.27%
	1	0.76	0.64	0.69	
SVM	0	0.78	0.86	0.82	77.27%
	1	0.76	0.64	0.69	

Table 6.1 Performance Metrics of the Model

B. User-Input Prediction

At the stage of predicting user inputs, a grayscale eye-tracking image is fed by a user via an input file path. Pixel-level features such as mean intensity, standard deviation, and active pixel count are extracted automatically from the image. Such features are fed to the trained ML model (say, Random Forest or SVM) for classifying the image. The model predicts whether the child is likely to be ASD (TS) or Non-ASD (TC). Prediction confidence or probability is also displayed for interpretability and decision-making support.

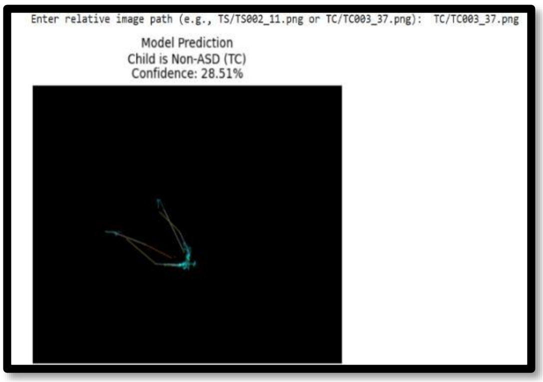

Fig 6.2 Prediction of Non-ASD

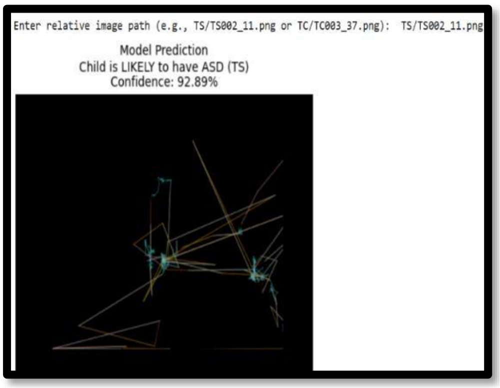

Fig 6.3 Prediction of ASD

VII. CONCLUSION AND FUTURE WORK

There may be future extensions of this study that would involve enriching a more sophisticated feature set by adding more detailed visual behavior metrics for fixation duration, saccade dynamics, and gaze transitions, which may provide deeper insights regarding patterns of attention in children with ASD.With the progress of deep learning methods, more specifically convolutional neural networks (CNNs), it gradually becomes possible to extract automatically complex spatial features from the scanning images and thus improve classification accuracy. It may help the model capture both where and how visual attention switches over time when temporal dynamics of eye-tracking data get added. Such a model could be a further improvement in the diagnostic accuracy. Finally, the validation of the model on a larger, more diverse population and naturalistic settings such as clinics or schools would ensure the generalization and practicality of the application. With these improvements, earlier detection of ASD would also be helped by a strong, scalable, and explainable approach to diagnostic aid development.

.

REFERENCES

[1] Z. Y. Zhao, J. M. Smith, and A. B. Lee, "Early Identification of Autism Spectrum Disorder Based on Machine Learning with Eye Tracking Data," J. Affect. Disord., vol. 358, pp. 326– 334, 2024.
[2] J. Wan, L. Wang, and F. Chen, "Gaze Fixation Metrics for Classifying Children with Autism Spectrum Disorder," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 27, no. 2, pp. 267–275, 2019.
[3] Y. Jiang, C. Luo, and M. Zhang, "Multimodal Detection of ASD Using Random Forests on Eye-Tracking and Behavioral Data," Comput. Biol. Med., vol. 114, pp. 103424, 2019.
[4] C. Liu, Y. Wu, and H. Wang, "Classification of Children with Autism Using Eye Movement Data and SVM," Autism Res., vol. 9, no. 6, pp. 653–662, 2016.
[5] A. Elbattah B. Molloy, and A. Salim, "Unsupervised Learning of Gaze-Based Features in ASD Detection,"J. Med. Syst., vol. 43, no. 9, pp. 254–262, 2019.
[6] H. Tsuchiya, Y. Saito, and R. Ogawa, "Age- Sensitive Gaze Detection for Autism Diagnosis Using Gazefinder," Front. Psychiatry, vol. 12,pp. 715214, 2021.
[7] S. Cilia, M. Gori, and F. Rinaldi, "Convolutional Neural Network for Eye-Movement-Based ASD Classification," IEEE Access, vol. 9, pp. 80361–80369, 2021.
[8]R. Kanhirakadavath and P. Chandran, "Motion-Based Gaze Feature Extraction for Autism Classification," Neurocomputing, vol. 509, pp. 400–410, 2022.
[9] X. Lin, D. Liu, and Z. Huang, "Hybrid Gaze and Demographic Features for ASD Detection Using Tree-Based Models," BMC Bioinformatics, vol. 23, pp. 94, 2022.
[10] H. Kong, Y. Park, and T. Lee, "Early Detection of Autism Using Eye Movement and Machine Learning," Pediatr. Res., vol. 93, pp. 356–364, 2023. .
[11] C. Fernández Núñez, J. J. Rodríguez, and A. Ledezma, "Real-Time CNN Gaze Preference Prediction for Autism Detection," Expert Syst. Appl., vol. 161, pp. 113701, 2020.
[12] P. Kasri, M. Abid, and R. Bennacer, "Vision Transformer and Mamba Hybrid for Eye-Based Autism Prediction," IEEE Trans. Cogn. Dev. Syst., vol. 17, no. 2, pp.

AUTHOR'S BIOGRAPHY

Full name	**DR.S.THARANI**
Science degree	Ph.D.
Academic rank	Assistant Professor
Institution	United College of Arts and Science

Full name	**M.KARTHIYAYINI**
Science degree	M.Sc. CS
Academic rank	Assistant Professor
Institution	United College of Arts and Science

Full name	**M.PRIYATHARSHINI**
Science degree	M.C.A
Academic rank	Assistant Professor
Institution	United College of Arts and Science

Full name	**A.YOGESWARI**
Science degree	M.Phil.
Academic rank	Assistant Professor
Institution	United College of Arts and Science

Full name	**V.VINODKANNA**
Science degree	M.Phil.
Academic rank	Assistant Professor
Institution	United College of Arts and Science